PAPER

Analysis Based on Moment Vector Equation for Interacting Identical Elements with Nonlinear Dynamics

Hideki SATOH[†], ***Member***

SUMMARY A method was developed for analyzing a system comprised of identical and indistinguishable elements with nonlinear dynamics. First, a moment vector equation (MVE) for the system was derived so as to avoid the curse of dimensionality by using the property that the elements are identical and indistinguishable. Next, an algorithm was developed to solve the MVE for deriving the moment vector in a steady state. It effectively uses eigen analysis on the basis of the property of the MVE. It can thus be used to clarify the structure of the solutions in the moment vector space and to derive multiple solutions by setting the initial value to the moment vector orthogonal to the solutions already obtained. Finally, the probability density function (pdf) for the state of the system was derived using the moment vectors in a steady state. Comparison of the pdfs thereby derived with those derived using numerical simulation showed that the method provided good approximations of the pdfs. Moreover, multiple solutions that are difficult to do using numerical simulation were derived.

key words: *MVE, nonlinear, many-body, eigen analysis*

1. Introduction

Various systems comprised of many small elements and their interactions abound in the natural world surrounding us. There are many examples of such many-body systems: globally coupled maps [1], ad hoc wireless networks [2], the Sherrington-Kirkpatrick model in spin-glass theory, coupled Josephson junction circuits, neural networks, and the galaxy in which we live. It is thus important to understand the properties of many-body systems. However, there is much difficulty in gaining such understanding.

Several reasons account for this difficulty. For example, simulating a system with a large number of elements is difficult because the number of calculations required for computing the interactions increases exponentially with the number of elements. The accumulated numerical error is also a serious problem. The error degrades accuracy, making the solution meaningless. It is particularly difficult to trace the change in the states of elements over a long period of time because of the accumulated numerical error.

One approach to overcoming these problems is to use a one-dimensional self-gravitating system comprised of identical mass sheets [3]. It is easier to numerically compute interactions by approximating an original three-dimensional system as a one-dimensional system. Another approach is to use Boltzmann equations to analytically obtain solutions on the basis of statistical mechanics [4], [5]. Although these two approaches can be used to obtain the macroscopic dynamics, they cannot always be applied to arbitrary many-body systems. Moreover, the accuracy of the solution is often limited. Nonlinear Fokker-Planck equations have been used to analyze more complex systems [6], [7], [8], but solving such equations has the same difficulty as solving other nonlinear equations - most nonlinear equations cannot be solved analytically.

Although various methods have been developed for deriving approximate analytical solutions, the properties for arbitrary nonlinear systems cannot be obtained. Although using a linearization approach at the equilibrium point makes it easy to obtain various properties around the equilibrium point, accurate solutions cannot be obtained far from the equilibrium point [9]. Perturbation analysis and asymptotic analysis may provide more accurate solutions to deterministic differential equations [10], [11], stochastic differential equations [12], and partial differential equations of a probability density function [13], but the accuracy does not always increase with the degree of the expansion. These analyses may sometimes require complex procedures, and they cannot be applied to every system. Moreover, their approximate solutions are rather complex, and multiple solutions cannot always be obtained. Thus, they are not always effective for analyzing nonlinear equations.

The moment vector equation (MVE) was developed to solve these problems [14]. It approximates a multi-dimensional nonlinear system as a linear vector equation by expanding the state space to a moment vector space. Various statistical properties, such as the mean, variance, covariance, and power spectrum, can be obtained with an algorithm based on the MVE, and the accuracy of the algorithm is increased by enlarging the dimension of the MVE. However, the dimension needed to obtain a given accuracy increases exponentially with the dimension of the state variable. This is referred to as the "curse of dimensionality" and is one of the most serious problems in function approximation and approximate analysis for high-dimensional systems.

A method was developed to avoid this curse as-

Manuscript received October 22, 2009.
Manuscript revised May 31, 2010.

[†]The author is with the Future University Hakodate, Hakodate-shi, 041-8655 Japan.

suming that the system is comprised of identical and indistinguishable elements. The method is presented in this paper, and its validity is demonstrated by showing not only that it provides good approximations but also that it finds solutions that are difficult to do using numerical simulation.

2. Analysis of Nonlinear System Using Moment Vector Equation

2.1 Moment Vector Equation for Nonlinear System

MVEs can be used to approximate an arbitrary multidimensional nonlinear system in the whole domain of definition [14]. Consider the following multidimensional discrete-time nonlinear system,

$$\boldsymbol{x}_{t+1} = \boldsymbol{f}(\boldsymbol{x}_t), \tag{1}$$

where $\boldsymbol{x}_t \stackrel{\text{def}}{=} (x_{1;t}, \cdots, x_{d_\mathrm{x};t})^\mathrm{T} \in \boldsymbol{\mathcal{D}}_\mathbf{x}$ is the state of dimension d_x, $\boldsymbol{f}(\cdot) \stackrel{\text{def}}{=} (f_1(\cdot), \cdots, f_{d_\mathrm{x}}(\cdot))^\mathrm{T}$ is a deterministic or stochastic function, subscript t denotes a discrete time, $\boldsymbol{\mathcal{D}}_\mathbf{x} \stackrel{\text{def}}{=} \{\boldsymbol{x}_t | x_{\min d} < x_{d;t} < x_{\max d}, 1 \le d \le d_\mathrm{x}\}$ is the domain of definition[†], and superscript T denotes a transposition.

Let $\{\psi_i(\cdot)\}$ be an orthonormal basis and $\psi_0(\cdot)$ be constant ψ_0 as defined in Appendix A. To derive the MVE for the nonlinear system in Eq. (1), the following assumption is introduced with respect to Eq. (1).

Assumption 1: We can expand $E[\psi_i(\boldsymbol{x}_{t+1})|\boldsymbol{x}_t]$ in a Fourier series:

$$E[\psi_i(\boldsymbol{x}_{t+1})|\boldsymbol{x}_t] = \sum_{j=0}^{N} a_{ij}\psi_j(\boldsymbol{x}_t) + \varepsilon_i(\boldsymbol{x}_t), \tag{2}$$

where $E[\cdot]$ is the mathematical expectation, $\varepsilon_i(\boldsymbol{x}_t)$ is the residual, and N is the degree of expansion. □

The above Fourier series converges in the usual sense; that is, $\varepsilon_i(\boldsymbol{x}_t)$ tends to zero as $N \to \infty$ if $E[\psi_i(\boldsymbol{x}_{t+1})|\boldsymbol{x}_t]$ satisfies the Dirichlet conditions [15]. In contrast, if $E[\psi_i(\boldsymbol{x}_{t+1})|\boldsymbol{x}_t]$ is discontinuous and $\psi_j(\boldsymbol{x}_t)$ is the trigonometric function defined in Eq. (A·7), the Gibbs phenomenon occurs. That is, $\varepsilon_i(\boldsymbol{x}_t)$ does not tend to zero in the vicinity of the discontinuity points no matter how large N is. This phenomenon can be avoided by selecting an appropriate basis [16]. We can thus make $\varepsilon_i(\boldsymbol{x}_t)$ sufficiently small by setting N to a sufficiently large value and selecting a basis suitable for the shape of $E[\psi_i(\boldsymbol{x}_{t+1})|\boldsymbol{x}_t]$.

Using Eq. (2), we can expand $E[\psi_i(\boldsymbol{x}_{t+1})]$:

$$\begin{aligned}
E[\psi_i(\boldsymbol{x}_{t+1})] &= \int \psi_i' p(\psi_i') \mathrm{d}\psi_i' \\
&= \int \psi_i' \int p(\boldsymbol{x}_t) p(\psi_i'|\boldsymbol{x}_t) \mathrm{d}\boldsymbol{x}_t \mathrm{d}\psi_i' \\
&= \int p(\boldsymbol{x}_t) \int \psi_i' p(\psi_i'|\boldsymbol{x}_t) \mathrm{d}\psi_i' \mathrm{d}\boldsymbol{x}_t \\
&= \int p(\boldsymbol{x}_t) E[\psi_i'|\boldsymbol{x}_t] \mathrm{d}\boldsymbol{x}_t \\
&= \int p(\boldsymbol{x}_t) (\sum_{j=0}^{N} a_{ij}\psi_j(\boldsymbol{x}_t) + \varepsilon_i(\boldsymbol{x}_t)) \mathrm{d}\boldsymbol{x}_t \\
&= \sum_{j=0}^{N} a_{ij} E[\psi_j(\boldsymbol{x}_t)] + E[\varepsilon_i(\boldsymbol{x}_t)],
\end{aligned} \tag{3}$$

where ψ_i' denotes $\psi_i(\boldsymbol{x}_{t+1})$ and $p(\cdot)$ denotes a probability density function (pdf). Consider an element that moves in accordance with Eq. (1). In this case, $p(\boldsymbol{x})$ is the pdf of the state, and $E[\cdot]$ is the expected value of the state. Now consider many elements each of which moves in accordance with Eq. (1) without interacting with the others. In this case, $p(\boldsymbol{x})$ is the pdf of the states, and $E[\cdot]$ is the mean of the states of all the elements.

When Eq. (1) is deterministic, a_{ij} is obtained using Eq. (A·2):

$$a_{ij} = \int_{\boldsymbol{\mathcal{D}}_\mathbf{x}} \psi_i(\boldsymbol{f}(\boldsymbol{x})) \psi_j^*(\boldsymbol{x}) \mathrm{d}\boldsymbol{x},$$

where superscript $*$ denotes a complex conjugate. We can assume that $\varepsilon_i(\boldsymbol{x}_t)$ is sufficiently small by setting N to a sufficiently large value and using an appropriate basis as mentioned above. Equation (3) thus can be reduced to

$$\boldsymbol{\rho}_{t+1} = A\boldsymbol{\rho}_t, \tag{4}$$

where $\boldsymbol{\psi}(\boldsymbol{x}_t) \stackrel{\text{def}}{=} (\psi_0(\boldsymbol{x}_t), \cdots, \psi_N(\boldsymbol{x}_t))^\mathrm{T}$, $\boldsymbol{\rho}_t \stackrel{\text{def}}{=} E[\boldsymbol{\psi}(\boldsymbol{x}_t)]$ is referred to as the moment vector, and A is the $(N+1) \times (N+1)$ matrix defined by

$$A \stackrel{\text{def}}{=} [a_{ij} | 0 \le i \le N, 0 \le j \le N].$$

Construction of the MVE in Eq. (4) means mapping $\boldsymbol{x}_t$ in a d_x-dimensional space to $\boldsymbol{\rho}_t$ in an $(N+1)$-dimensional space in which the nonlinear equation in Eq. (1) is approximately expressed by the linear equation in Eq. (4). Using Eq. (4), we can derive not only the expected value of $\psi_i(\boldsymbol{x}_t)$ but also the statistical properties such as the mean, variance, covariance, and power spectrum of $\boldsymbol{x}_t$. Because we can make $\varepsilon_i(\boldsymbol{x}_t)$ sufficiently small by setting N to a sufficiently large value, the accuracies of the statistical properties can be improved [14]. An example of the effect of N on the accuracy will be shown in Fig. 3 in Sect. 4.

Let N_d be the degree of expansion for x_d; it defines the accuracy of the statistical properties with respect

[†]When the domain of definition is infinite, an MVE can be derived using Hermite polynomials as an orthonormal basis $\{\psi_i\}$.

to x_d. Let $N_1,\cdots,N_{d_\mathrm{x}}$ be certain values; that is, an accuracy is set for $x_1,\cdots, x_{d_\mathrm{x}}$. We can then see from Eq. (A·6) that the degree of expansion for $\boldsymbol{x}$, N, increases geometrically with d_x. Therefore, we cannot set sufficient accuracy for $x_1,\cdots, x_{d_\mathrm{x}}$ if d_x is large [14]. This is the "curse of dimensionality," the most serious problem when the MVE is used for analyzing high-dimensional systems. A method was developed to avoid this curse assuming that the system can be approximated as a many-body system with identical and indistinguishable nonlinear elements. It is presented in Sect. 3.

2.2 Average of State and Moment Vector

Let us expand ${x_d}^n$ in a Fourier series as

$$ {x_d}^n = \sum_{j=0}^{N} X_{(n)dj}\psi_j(\boldsymbol{x}), \tag{5} $$

where $X_{(n)dj} = \int {x_d}^n \psi_j^*(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$. Let $X_{(n)}$ be the $d_\mathrm{x} \times (N+1)$ matrix defined by

$$ X_{(n)} \stackrel{\mathrm{def}}{=} [X_{(n)dj} | 1 \le d \le d_\mathrm{x}, 0 \le j \le N]. $$

Taking the expectation of Eq. (5) and using the definition of $\boldsymbol{\rho}$, we obtain

$$ (E[{x_1}^n], \cdots, E[{x_{d_\mathrm{x}}}^n])^\mathrm{T} = X_{(n)}\boldsymbol{\rho}. \tag{6} $$

Let $\langle y(t)\rangle$ be the infinite-time average of discrete-time variable $y(t)$ defined by

$$ \langle y(t)\rangle \stackrel{\mathrm{def}}{=} \lim_{T\to\infty} T^{-1}\sum_{\tau=0}^{T-1} y(t+\tau). $$

The following assumption is introduced for convenience of analysis to derive the infinite-time average of state $\boldsymbol{x}_t$ and that of moment vector $\boldsymbol{\rho}_t$.

Assumption 2: $\boldsymbol{\rho}_t$ does not diverge for $t \to \infty$. □

From Eq. (6), the infinite-time average of $(E[{x_{1;t}}^n], \cdots, E[{x_{d_\mathrm{x};t}}^n])^\mathrm{T}$ in a steady state† is obtained:

$$ \lim_{t\to\infty} \langle (E[{x_{1;t}}^n], \cdots, E[{x_{d_\mathrm{x};t}}^n])^\mathrm{T}\rangle = X_{(n)} \lim_{t\to\infty} \langle \boldsymbol{\rho}_t\rangle. $$

Let λ_i be the ith eigenvalue of matrix A, $\boldsymbol{e}_i$ be the ith eigenvector of matrix A $(0 \le i \le N)$, and λ_i and $\boldsymbol{e}_i$ be arranged by $|\lambda_i|$ so that $\lambda_0 = 1$ and $e_{00} \ne 0$. Here, $\forall|\lambda_i| \le 1$ holds from Assumption 2, and $\psi_0(\cdot)$ is a constant from the definition of $\{\psi_i(\cdot)\}$. Thus, there is at least one eigenvalue and eigenvector of matrix A such that $\lambda_i = 1$ and $e_{i0} \ne 0$. Because Eq. (4) is linear, $\langle\boldsymbol{\rho}_\infty\rangle$ is equal to equilibrium point $\boldsymbol{\rho}_\mathrm{e}$ such that $\boldsymbol{\rho}_\mathrm{e} = A\boldsymbol{\rho}_\mathrm{e}$ even if $\boldsymbol{\rho}_t$ oscillates at $t = \infty$ [14]. Therefore, $\langle\boldsymbol{\rho}_\infty\rangle$ is obtained by modifying the norm of $\boldsymbol{e}_0$ so that the first element in $\langle\boldsymbol{\rho}_\infty\rangle$ equals ψ_0.

$$ \begin{aligned} \langle\boldsymbol{\rho}_\infty\rangle &= \boldsymbol{\rho}_\mathrm{e} \\ &= (\psi_0/e_{00})\boldsymbol{e}_0. \end{aligned} \tag{7} $$

The above equation, based on the equilibrium point of Eq. (4), contains information about the statistical properties not only when $\boldsymbol{x}_t$ converges but also when $\boldsymbol{x}_t$ oscillates. A method for deriving the pdf of $\boldsymbol{x}_t$ will be presented in the next section.

2.3 Probability Density Function of State Variable

The pdf of $\boldsymbol{x}_t$ of the nonlinear system in Eq. (1) in a steady state is derived using $\langle\boldsymbol{\rho}_\infty\rangle$ in the same manner as the other statistical properties such as mean, variance, and power spectrum [14]. Consider N_Res appropriate pdfs $\hat{p}_1(\boldsymbol{x}),\cdots,\hat{p}_{N_\mathrm{Res}}(\boldsymbol{x})$ and weights $w_1,\cdots,w_{N_\mathrm{Res}}$. A pdf of $\boldsymbol{x}$, $p(\boldsymbol{x})$, can be expressed as

$$ p(\boldsymbol{x}) = \sum_{j=1}^{N_\mathrm{Res}} w_j\hat{p}_j(\boldsymbol{x}). \tag{8} $$

$E[\psi_i(\boldsymbol{x})]$ can be expressed using the above equation as

$$ \begin{aligned} E[\psi_i(\boldsymbol{x})] &= \int \psi_i(\boldsymbol{x})p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ &= \sum_{j=1}^{N_\mathrm{Res}} w_j \int \psi_i(\boldsymbol{x})\hat{p}_j(\boldsymbol{x})\mathrm{d}\boldsymbol{x}. \end{aligned} $$

The ith element of the moment vector at time t is derived as

$$ \begin{aligned} \rho_{i;t} &\stackrel{\mathrm{def}}{=} E[\psi_i(\boldsymbol{x}_t)] \\ &= \sum_{j=1}^{N_\mathrm{Res}} w_{j;t} \int \psi_i(\boldsymbol{x})\hat{p}_j(\boldsymbol{x})\mathrm{d}\boldsymbol{x}. \end{aligned} $$

We thus obtain

$$ \langle\boldsymbol{\rho}_\infty\rangle = W_\psi\langle\boldsymbol{w}_\infty\rangle, \tag{9} $$

where W_ψ is an $(N+1) \times N_\mathrm{Res}$ matrix defined by $[\int \psi_i(\boldsymbol{x})\hat{p}_j(\boldsymbol{x})\mathrm{d}\boldsymbol{x} | 0 \le i \le N, 1 \le j \le N_\mathrm{Res}]$ and $\boldsymbol{w}_t \stackrel{\mathrm{def}}{=} (w_{1;t}, \cdots, w_{N_\mathrm{Res};t})^\mathrm{T}$.

Let $\delta(\boldsymbol{x})$ be the delta function and $\hat{p}_j(\boldsymbol{x})$ be

$$ \hat{p}_j(\boldsymbol{x}) \stackrel{\mathrm{def}}{=} \delta(\boldsymbol{x} - \hat{\boldsymbol{x}}_j). \tag{10} $$

It is obvious from Eq. (8) that $w_1,\cdots,w_{N_\mathrm{Res}}$ are the values of the probability function when $\boldsymbol{x}$ is discretized to $\hat{\boldsymbol{x}}_1, \cdots, \hat{\boldsymbol{x}}_{N_\mathrm{Res}}$. Thus, normalizing and interpolating $\langle w_{1;\infty}\rangle, \cdots, \langle w_{N_\mathrm{Res};\infty}\rangle$ yields the pdf of $\boldsymbol{x}$ in Eq. (1) in a steady state.

The value of $\langle\boldsymbol{w}_\infty\rangle$ can be obtained as shown below. Using $\hat{p}_j(\boldsymbol{x})$ defined in Eq. (10), we obtain

$$ \begin{aligned} \int \psi_i(\boldsymbol{x})\hat{p}_j(\boldsymbol{x})\mathrm{d}\boldsymbol{x} &= \int \psi_i(\boldsymbol{x})\delta(\boldsymbol{x} - \hat{\boldsymbol{x}}_j)\mathrm{d}\boldsymbol{x} \\ &= \psi_i(\hat{\boldsymbol{x}}_j). \end{aligned} $$

†The system is said in this paper to be in a steady state if the time average of the characteristics does not change with time.

From this equation, W_ψ is expressed as

$$W_\psi = [\psi_i(\hat{\boldsymbol{x}}_j)|0 \leq i \leq N, 1 \leq j \leq N_{\mathrm{Res}}].$$

We can then obtain $\langle \boldsymbol{w}_\infty \rangle$ by modifying Eq. (9):

$$\langle \boldsymbol{w}_\infty \rangle = {W_\psi}^- \langle \boldsymbol{\rho}_\infty \rangle. \tag{11}$$

Here, ${W_\psi}^-$ is the generalized inverse matrix of W_ψ.

3. MVEs for Many-body System

MVEs for analyzing the interactions between a large number of elements are presented and used to analyze the statistical properties of a many-body system.

3.1 MVE for Linear Interaction

Consider a discrete-time system with L elements in which the following assumption holds.

Assumption 3: The elements are identical and indistinguishable, and the interactions between the elements are determined by only the states of the elements. □

Assumption 3 basically holds in many systems with a large number of elements such as self-gravitating many-body systems.

Let $\boldsymbol{x}_{(\ell)t}$ be the state of the ℓth element at time t and $L^{-1}\hat{\boldsymbol{h}}(\boldsymbol{x}_{(\ell)t}, \boldsymbol{x}_{(k)t})$ be the interaction from the kth element to the ℓth one. Then, we can describe the system as

$$\boldsymbol{x}_{(\ell)t+1} = \boldsymbol{x}_{(\ell)t} + \sum_{k=1}^{L} L^{-1}\hat{\boldsymbol{h}}(\boldsymbol{x}_{(\ell)t}, \boldsymbol{x}_{(k)t}).$$

When the interactions are linear, as defined by

$$\hat{\boldsymbol{h}}(\boldsymbol{x}_{(\ell)t}, \boldsymbol{x}_{(k)t}) \stackrel{\mathrm{def}}{=} c(\boldsymbol{x}_{(\ell)t} - \boldsymbol{x}_{(k)t}),$$

the above system is expressed as

$$\boldsymbol{x}_{(\ell)t+1} = \boldsymbol{x}_{(\ell)t} + \sum_{k=1}^{L} L^{-1}c(\boldsymbol{x}_{(\ell)t} - \boldsymbol{x}_{(k)t}). \tag{12}$$

Note that L^{-1} is introduced to simplify equation expansion. The above equation is a globally coupled map, and such a map with chaotic elements has many attractive features from the viewpoint of biological information processing and engineering applications [1].

Equation (12) is reduced to

$$\boldsymbol{x}_{(\ell)t+1} = \boldsymbol{x}_{(\ell)t} + c(\boldsymbol{x}_{(\ell)t} - E[\boldsymbol{x}_{(k)t}|_{1\leq k\leq L}]). \tag{13}$$

Here, if $c > 0$, $\boldsymbol{x}_{(\ell)t+1}$ diverges from $E[\boldsymbol{x}_{(k)t}|_{1\leq k\leq L}]$, and if $c < 0$, $\boldsymbol{x}_{(\ell)t+1}$ approaches $E[\boldsymbol{x}_{(k)t}|_{1\leq k\leq L}]$. Because we assume that the elements are identical and indistinguishable in Assumption 3, suffix (ℓ) is omitted, and $E[\boldsymbol{x}_{(k)t}|_{1\leq k\leq L}]$ is abbreviated to $E[\boldsymbol{x}_t]$. The following equation is thus used instead of Eq. (13).

$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t + c(\boldsymbol{x}_t - E[\boldsymbol{x}_t]). \tag{14}$$

The above equation is modified, resulting in

$$\boldsymbol{x}_{t+1} = c_1\boldsymbol{x}_t + c_2E[\boldsymbol{x}_t], \tag{15}$$

where $c_1 = (1 + c)$ and $c_2 = -c$. In the same manner as in Sect. 2.1, $E[\psi_i(\boldsymbol{x}_{t+1})]$ for the above equation is expanded:

$$E[\psi_i(\boldsymbol{x}_{t+1})] = \sum_{j=0}^{N} b_{ij}E[\psi_j(\boldsymbol{x}_t)],$$
$$b_{ij} \stackrel{\mathrm{def}}{=} \psi_i(c_2E[\boldsymbol{x}_t]) \int \psi_i(c_1\boldsymbol{x}_t)\psi_j^*(\boldsymbol{x}_t)\mathrm{d}\boldsymbol{x}_t.$$

Substituting $\boldsymbol{\rho} \stackrel{\mathrm{def}}{=} E[\boldsymbol{\psi}(\boldsymbol{x})]$ and Eq. (6) into the above two equations yields the MVE of Eq. (15):

$$\boldsymbol{\rho}_{t+1} = B(\boldsymbol{\rho}_t)\boldsymbol{\rho}_t, \tag{16}$$

where $B(\boldsymbol{\rho}_t) \stackrel{\mathrm{def}}{=} [b_{ij}(\boldsymbol{\rho}_t)|0 \leq i \leq N, 0 \leq j \leq N]$ is an $(N + 1) \times (N + 1)$ matrix. Matrix element $b_{ij}(\boldsymbol{\rho}_t)$ is reduced to

$$b_{ij}(\boldsymbol{\rho}_t) \stackrel{\mathrm{def}}{=} \psi_i(c_2X_{(1)}\boldsymbol{\rho}_t)\xi_{ij},$$

and $\xi_{ij} \stackrel{\mathrm{def}}{=} \int \psi_i(c_1\boldsymbol{x}_t)\psi_j^*(\boldsymbol{x}_t)\mathrm{d}\boldsymbol{x}_t$.

3.2 MVE for Linear Interaction of Elements with Nonlinear Dynamics

Consider a discrete-time system with L identical and indistinguishable elements described by

$$\boldsymbol{x}_{(\mathrm{A})t+1} = \boldsymbol{g}_{(\mathrm{A})}(\boldsymbol{f}_{(A)}(\boldsymbol{x}_{(\mathrm{A})t})), \tag{17}$$

where $\boldsymbol{x}_{(A)} \stackrel{\mathrm{def}}{=} ({\boldsymbol{x}_{(1)}}^{\mathrm{T}}, \cdots, {\boldsymbol{x}_{(L)}}^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{f}_{(A)}(\boldsymbol{x}_{(A)}) \stackrel{\mathrm{def}}{=} (\boldsymbol{f}(\boldsymbol{x}_{(1)})^{\mathrm{T}}, \cdots, \boldsymbol{f}(\boldsymbol{x}_{(L)})^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{f}$ denotes the nonlinear dynamics of each element described in Eq. (1), $\boldsymbol{g}_{(\mathrm{A})}(\boldsymbol{x}_{(A)}) \stackrel{\mathrm{def}}{=} (\boldsymbol{g}_{(1)}(\boldsymbol{x}_{(A)})^{\mathrm{T}}, \cdots, \boldsymbol{g}_{(L)}(\boldsymbol{x}_{(A)})^{\mathrm{T}})^{\mathrm{T}}$, and $\boldsymbol{g}_{(\ell)}(\boldsymbol{x}_{(A)})$ denotes the linear interaction defined by the right-hand side of Eq. (12), i.e.,

$$\boldsymbol{g}_{(\ell)}(\boldsymbol{x}_{(A)}) \stackrel{\mathrm{def}}{=} \boldsymbol{x}_{(\ell)t} + \sum_{k=1}^{L} L^{-1}c(\boldsymbol{x}_{(\ell)t} - \boldsymbol{x}_{(k)t}). \tag{18}$$

The system is shown in Fig. 1.

Equation (17) is divided into two parts:

$$\boldsymbol{x}'_{(A)} = \boldsymbol{f}_{(A)}(\boldsymbol{x}_{(\mathrm{A})t}),$$
$$\boldsymbol{x}_{(\mathrm{A})t+1} = \boldsymbol{g}_{(\mathrm{A})}(\boldsymbol{x}'_{(A)}).$$

In the same manner as reducing Eq. (12) to Eq. (14), suffix (ℓ) can be omitted, and the above equations can be reduced to

$$\boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x}_t),$$
$$\boldsymbol{x}_{t+1} = \boldsymbol{x}' + c(\boldsymbol{x}' - E[\boldsymbol{x}']).$$

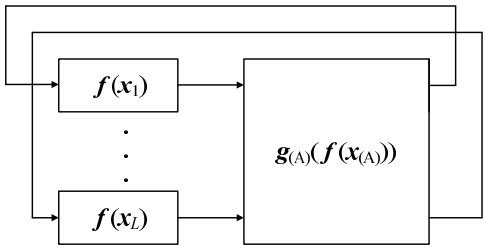

Fig. 1 Many-body system: L elements with nonlinear dynamics and their interactions.

Let $\boldsymbol{\rho}'$ be $E[\psi(\boldsymbol{x}')]$. The MVEs of the above equations can be expressed as

$$\begin{aligned}\boldsymbol{\rho}' &= A\boldsymbol{\rho}_t,\\ \boldsymbol{\rho}_{t+1} &= B(\boldsymbol{\rho}')\boldsymbol{\rho}',\end{aligned}$$

in the same manner as for Eqs. (4) and (16). Thus, the MVE of Eq. (17) is expressed as

$$\boldsymbol{\rho}_{t+1} = B(A\boldsymbol{\rho}_t)A\boldsymbol{\rho}_t. \tag{19}$$

We can also derive the MVE of Eq. (17) using only the procedure in Sect. 2. However, its matrix size is $(N+1)^L - 1$ because the basis for the system has to be the direct product of the basis for the elements [14]. This is the curse of dimensionality, one of the most serious problems in function approximation and analysis of large-scale or high-dimensional nonlinear systems. So the MVE of Eq. (17) that can be used in practice cannot be derived using only the procedure in Sect. 2. The procedure described above was developed to derive an MVE that avoids the curse of dimensionality assuming that the elements are identical and indistinguishable, as in Assumption 3. This MVE is nonlinear, as shown in Eq. (19). Thus, we cannot use the various methods based on a linear MVE for deriving the statistical properties [14] although we can avoid the curse of dimensionality. Two algorithms have been developed to analyze the system on the basis of the nonlinear MVE in Eq. (19), and they will be presented in the next section.

3.3 Probability Density Function of State Variable

3.3.1 Relationship between Infinite-time Average and Equilibrium Point

As described in Sect. 2.3, the pdf of $\boldsymbol{x}_t$ of the nonlinear system in Eq. (1) is derived using Eq. (11) and $\langle\boldsymbol{\rho}_\infty\rangle$ obtained from the MVE of Eq. (4). Because Eq. (4) is linear, $\langle\boldsymbol{\rho}_\infty\rangle$ is equal to equilibrium point $\boldsymbol{\rho}_\mathrm{e}$ of the MVE of Eq. (4), so $\langle\boldsymbol{\rho}_\infty\rangle$ can be derived by eigen analysis of matrix A [14].

The pdf of $\boldsymbol{x}_t$ in Eq. (17) is derived using Eq. (11) in the same manner as for Eq. (1) if $\langle\boldsymbol{\rho}_\infty\rangle$ for Eq. (19) is obtained. However, Eq. (19) is nonlinear, so $\langle\boldsymbol{\rho}_\infty\rangle$ is not equal to equilibrium point $\boldsymbol{\rho}_\mathrm{e}$. The set of equilibrium points is a subset of $\langle\boldsymbol{\rho}_\infty\rangle$[†]. Specifically, the pdf when $E[\boldsymbol{x}_t]$ oscillates cannot be obtained from equilibrium point $\boldsymbol{\rho}_\mathrm{e}$ for Eq. (19). It is thus very difficult to derive $\langle\boldsymbol{\rho}_\infty\rangle$. Two algorithms that use eigen analysis to derive $\langle\boldsymbol{\rho}_\infty\rangle$ of Eq. (19) effectively are presented below.

3.3.2 Pdf for Convergent Moment Vector Sequence

Under Assumption 2, there is not only the case where $\boldsymbol{\rho}_t$ converges to a constant but also the case where $\boldsymbol{\rho}_t$ oscillates with period T_ρ[††]. Although it is difficult to directly analyze oscillation sequence $\boldsymbol{\rho}_t, \boldsymbol{\rho}_{t+1}, \boldsymbol{\rho}_{t+2}, \cdots$ with period T_ρ, sequence $\boldsymbol{\rho}_t, \boldsymbol{\rho}_{t+T_\rho}, \boldsymbol{\rho}_{t+2T_\rho}, \cdots$ converges to a constant, and the analysis of the sequence is easier than that of the oscillation sequence. Let us thus consider the following equation instead of Eq. (19):

$$\boldsymbol{\rho}_{t+T_\rho} = A_\mathrm{comb}(\boldsymbol{\rho}_t, T_\rho)\boldsymbol{\rho}_t, \tag{20}$$

where $T_\rho = 1, 2, \cdots$, $A_\mathrm{comb}(\boldsymbol{\rho}_t, T_\rho)$ is defined by the recursive function of Eq. (19) as

$$A_\mathrm{comb}(\boldsymbol{\rho}, T_\rho) \stackrel{\mathrm{def}}{=} \begin{cases} B(A\boldsymbol{\rho})A & \text{for } T_\rho = 1\\ A_\mathrm{comb}(A_\mathrm{comb}(\boldsymbol{\rho}, T_\rho - 1)\boldsymbol{\rho}, 1) & \\ \quad A_\mathrm{comb}(\boldsymbol{\rho}, T_\rho - 1) & \text{for } T_\rho \geq 2. \end{cases}$$

When $\boldsymbol{\rho}_t$ converges to a constant, the equilibrium point of Eq. (20) with $T_\rho = 1$ is equal to $\langle\boldsymbol{\rho}_\infty\rangle$ in Eq. (19) in the same manner as in Sect. 2.2. When $\boldsymbol{\rho}_t$ oscillates, we can estimate the properties of $\boldsymbol{x}_t$ in Eq. (17) by investigating equilibrium point $\boldsymbol{\rho}_\mathrm{e}$ of Eq. (20) for various values of T_ρ in the same manner as in the Poincare map used for analyzing nonlinear equations [17].

The equilibrium point, $\boldsymbol{\rho}_\mathrm{e}$, of Eq. (20) can be derived as a solution to

$$\boldsymbol{\rho} = A_\mathrm{comb}(\boldsymbol{\rho}, T_\rho)\boldsymbol{\rho}. \tag{21}$$

The above equation is nonlinear, and the dimension of $\boldsymbol{\rho}$ is high. It thus takes a long time to find $\boldsymbol{\rho}_\mathrm{e}$ if we use a conventional method such as the steepest descent method or the Newton method. However, if $\boldsymbol{\rho}$ in matrix $A_\mathrm{comb}(\boldsymbol{\rho}, T_\rho)$ in Eq. (21) is fixed, Eq. (21) becomes a linear equation, so eigen analysis may be effective. Algorithm 1 was developed using this property of Eq. (21). It uses eigen analysis to effectively derive equilibrium point $\boldsymbol{\rho}_\mathrm{e}$ of sequence $\boldsymbol{\rho}_t, \boldsymbol{\rho}_{t+T_\rho}, \boldsymbol{\rho}_{t+2T_\rho}, \cdots$ within only a few iterations. The number of iterations needed will be shown in Sect. 4.

[†]There is at least one equilibrium point $\boldsymbol{\rho}_\mathrm{e}$ in Eq. (19) although Eq. (19) may not be stable at $\boldsymbol{\rho}_\mathrm{e}$ ($\boldsymbol{\rho}_\mathrm{e}$ may be an unstable equilibrium point) even if $\boldsymbol{\rho}_t$ oscillates at $t = \infty$ because Eq. (19) does not diverge under Assumption 2.

[††]Assumption 2 often holds even if the original system does not have a steady state because the MVE of Eq. (19) is an approximation. Systems with an MVE for which Assumption 2 holds are considered in this paper.

Algorithm 1:

(1-1) Set $\boldsymbol{\rho}_0$ and T_ρ.

(1-2) $t = 0$.

(1-3) Compute $A_{\mathrm{comb}}(\boldsymbol{\rho}_t, T_\rho)$.

(1-4) Derive eigenvector of $A_{\mathrm{comb}}(\boldsymbol{\rho}_t, T_\rho)$ with $\lambda_0 = 1$ and $[\boldsymbol{e}_0]_0 \neq 0$, and set the right-hand side of Eq. (7) to $\boldsymbol{\rho}_{t+1}$.

(1-5) If $\|\boldsymbol{\rho}_{t+1} - \boldsymbol{\rho}_t\| < \varepsilon$, set $\boldsymbol{\rho}_{t+1}$ to $\boldsymbol{\rho}_{\mathrm{e}}$ and go to Step (1-7). Otherwise, go to Step (1-6).

(1-6) Set $t = t + 1$ and go back to Step (1-3).

(1-7) Set $\boldsymbol{\rho}_{\mathrm{e}}$ to $\langle \boldsymbol{\rho}_\infty \rangle$ and derive the pdf using Eq. (11).

Here, ε is an allowable error to finish the algorithm. With Algorithm 1, a pdf of sequence $\boldsymbol{x}_t$, $\boldsymbol{x}_{t+T_\rho}$, $\boldsymbol{x}_{t+2T_\rho}$, $\cdots$ when sequence $\boldsymbol{\rho}_t$, $\boldsymbol{\rho}_{t+T_\rho}$, $\boldsymbol{\rho}_{t+2T_\rho}$, $\cdots$ converges is obtained.

3.3.3 Pdf for Oscillating Moment Vector Sequence

When sequence $\boldsymbol{\rho}_t$, $\boldsymbol{\rho}_{t+1}$, $\boldsymbol{\rho}_{t+2}$, $\cdots$ does not converge, that is, $E[\boldsymbol{x}_t]$ oscillates, the pdf of $\boldsymbol{x}_t$ cannot be derived using Algorithm 1. Therefore, it was expanded into Algorithm 2. Let $\boldsymbol{\rho}_{\mathrm{e}\tau}$ be the τth equilibrium point for Eq. (20) and N_{e} be the number of equilibrium points. When $\boldsymbol{\rho}_t$ oscillates with period $\hat{T}_\rho$, there are $\hat{T}_\rho$ equilibrium points in Eq. (20) if we set $T_\rho = \hat{T}_\rho$. Thus, if $N_{\mathrm{e}} = T_\rho$, $\langle \boldsymbol{\rho}_\infty \rangle$ is obtained as the mean of the equilibrium points of Eq. (20) as follows.

$$\langle \boldsymbol{\rho}_\infty \rangle = {N_{\mathrm{e}}}^{-1} \sum_{\tau=1}^{N_{\mathrm{e}}} \boldsymbol{\rho}_{\mathrm{e}\tau}. \tag{22}$$

The pdf when $\boldsymbol{\rho}_t$ oscillates is obtained using $\langle \boldsymbol{\rho}_\infty \rangle$ derived using the above equation and Eq. (11).

Multiple equilibrium points, $\boldsymbol{\rho}_{\mathrm{e}1}$, $\boldsymbol{\rho}_{\mathrm{e}2}$, $\cdots$, are obtained by executing Algorithm 1 for different initial values. However, a way for selecting the initial values for the nonlinear programming has not been established yet, so we cannot always derive all the solutions of a nonlinear equation. Algorithm 2 was developed for deriving as many equilibrium points of Eq. (20) as possible.

Algorithm 2:

(2-1) Set T_ρ and $\boldsymbol{\rho}_0 = (1, 0, \cdots, 0)^{\mathrm{T}}$ and perform Steps (1-2) through (1-6) in Algorithm 1. Let the solution obtained be $\boldsymbol{\rho}_{\mathrm{sol}}$.

(2-2) Set $i = 1$.

(2-3) Set the ith eigenvector of $A_{\mathrm{comb}}(\boldsymbol{\rho}_{\mathrm{sol}}, T_\rho)$ to $\boldsymbol{\rho}_0$ and perform Steps (1-2) through (1-6) in Algorithm 1. Let the solution obtained be $\boldsymbol{\rho}_{\mathrm{sol}(i)}$.

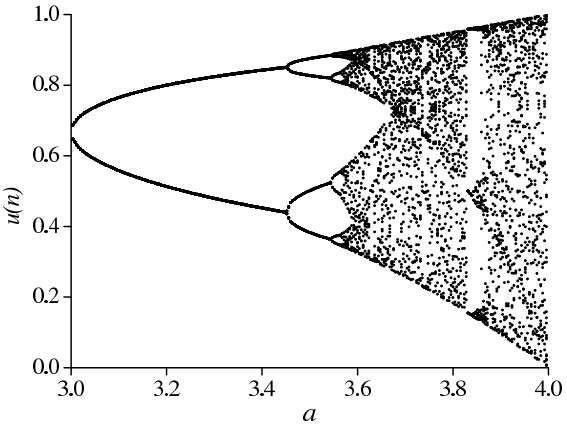

Fig. 2 Bifurcation diagram of logistic map [17].

(2-4) If T_ρ different solutions are found in $\{\boldsymbol{\rho}_{\mathrm{sol}}, \boldsymbol{\rho}_{\mathrm{sol}(1)}, \cdots\}$, set them to $\boldsymbol{\rho}_{\mathrm{e}1}$, $\cdots$, $\boldsymbol{\rho}_{\mathrm{e}T_\rho}$ and go to Step (2-5). If $i = N_{\mathrm{eigen}}$, stop. Otherwise, set $i = i + 1$ and go back to Step (2-3).

(2-5) Set $\langle \boldsymbol{\rho}_\infty \rangle$ using Eq. (22) and derive the pdf using Eq. (11).

Here, N_{eigen} is the maximum number of eigenvectors used as the initial values.

In Algorithm 2, an equilibrium point, $\boldsymbol{\rho}_{\mathrm{sol}}$, is first derived in Step (2-1), and the eigenvectors of $A_{\mathrm{comb}}(\boldsymbol{\rho}_{\mathrm{sol}}, T_\rho)$ are used as the initial values in Step (2-3). Although it has not been proved that Algorithm 2 can always derive all the equilibrium points for Eq. (20), using the initial values based on the eigenvectors is effective. This is because the eigenvectors are orthogonal to not only $\boldsymbol{\rho}_{\mathrm{sol}}$ but also each other and thus are the moment vectors most different from $\boldsymbol{\rho}_{\mathrm{sol}}$. Performing Algorithm 2 for various values of T_ρ should result in various pdfs being obtained more easily than by performing numerical simulation. The performance of Algorithm 2 will be described in Sect. 4.

4. Performance Evaluation

Consider the system defined by Eqs. (17) and (18). Here, $d_{\mathrm{x}} = 1$; f, which expresses the nonlinear dynamics of each element, is the logistic map [17] described by

$$f(x) = ax(1 - x), \tag{23}$$

and $0 < a < 4$ is the parameter of the logistic map. The system is the same as a globally coupled map [1].

Figure 2 shows the well-known bifurcation diagram of a logistic map. As we can see from the diagram, the logistic map shows complicated behavior in response to changes in parameter a. Analysis of the system described in Eqs. (17) and (18) is thus a good way to evaluate the performance of Algorithm 1 if the intrinsic

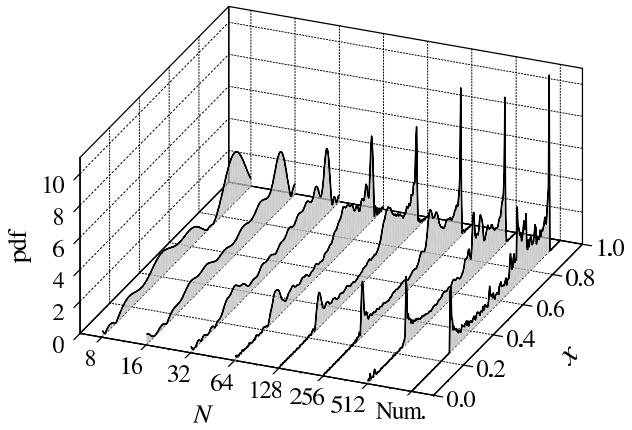

Fig. 3 Effect of N on accuracy of Algorithm 1 ($a = 3.7$ and $c = 0$).

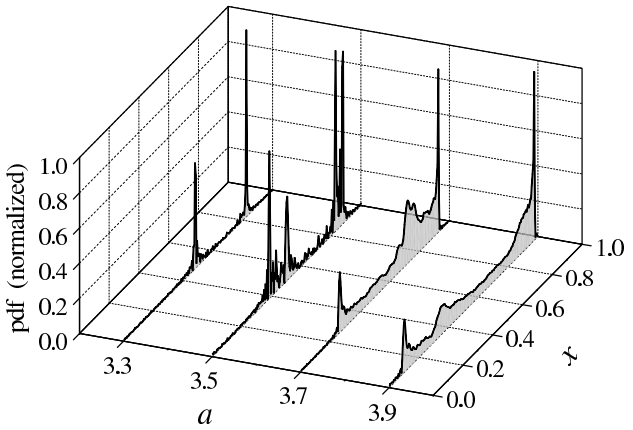

Fig. 4 Pdfs for various values of a with Algorithm 1 ($c = 0$ and $N = 256$).

motion of each element obeys the logistic map. Therefore, the logistic map was used to express the nonlinear dynamics of each element. The pdfs of the system in a steady state were derived using Algorithm 1 and numerical simulation. The parameters of Algorithm 1 were set as $\boldsymbol{\rho}_0 = (1, 0, \cdots, 0)^{\mathrm{T}}$, which means that the initial values of the state, x_0, are distributed uniformly, $T_\rho = 1$, and $\varepsilon = 10^{-3}$. The number of elements, L, was set to 128. Note that suffix (ℓ) for identifying each element is omitted and $x_{(\ell)t}$ is abbreviated as x_t, as mentioned in Sect. 3.1. Thus, x_t represents $x_{(1)}, \cdots, x_{(L)}$.

First, let us consider the relationship between the accuracy of Algorithm 1 and N, the degree of expansion. The pdfs were evaluated for $a = 3.7$ and $c = 0$ using Algorithm 1 and numerical simulation. Here, $c = 0$ means that there is no interaction between the elements, and the value was used because the precise pdf was obtained by numerical simulation, enabling us to compare the results obtained using Algorithm 1 with those obtained using numerical simulation.

Figure 3 shows the pdfs obtained using Algorithm 1 for various values of N and numerical simulation. The "Num." on the abscissa indicates the result obtained by numerical simulation. As shown in Fig. 3, the shapes of the pdfs obtained using Algorithm 1 approached that obtained using numerical simulation as N increased. That is, the accuracy of Algorithm 1 increased with N.

The pdfs were also evaluated for various values of a and $c = 0$ using Algorithm 1 and numerical simulation, as shown in Figs. 4 and 5. Here, N was set to 256, and the pdfs were normalized on the basis of their peak values for convenience of comparison. The pdfs obtained using Algorithm 1 are very similar to those obtained using numerical simulation, and they are consistent with the bifurcation diagram in Fig. 2. Thus, Algorithm 1 can be used to derive accurate pdfs for various values of a.

Next, pdfs were examined for $a = 3.7$ and various

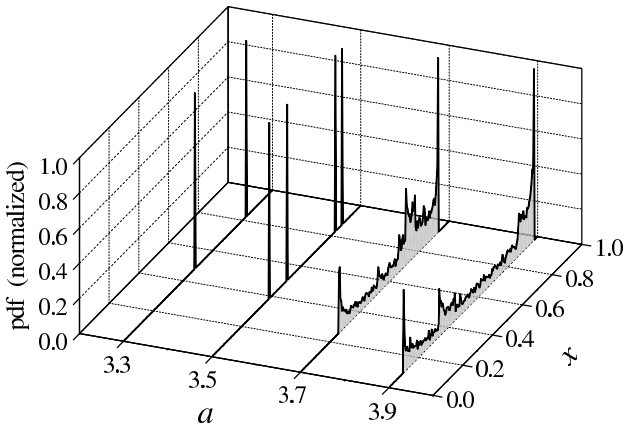

Fig. 5 Pdfs for various values of a by numerical simulation ($c = 0$).

values of c to investigate whether Algorithm 1 works when the elements interact with each other. Figure 6 shows the pdfs obtained using Algorithm 1 ($N = 256$). When the value of c was near zero ($c = -0.025$), that is, the interactions were small, the elements experienced chaotic oscillations, and x_t was distributed widely in the domain of definition. The pdf was almost equal to that when there were no interactions (See Figs. 4 and 5 for $a = 3.7$). As c diverged from zero, that is, the interactions increased, the chaotic oscillation decreased, and x took restricted values. These properties obtained using Algorithm 1 are almost the same as those obtained using numerical simulation, which are shown in Fig. 7. Moreover, almost the same pdfs between Algorithm 1 and numerical simulation were obtained for $c = -0.05$ and various values of a, as shown in Figs. 8 and 9. Algorithm 1 can thus be used to analyze a system with interactions between the elements.

Finally, it is shown that Algorithm 2 can be used to derive solutions that are very difficult to obtain by

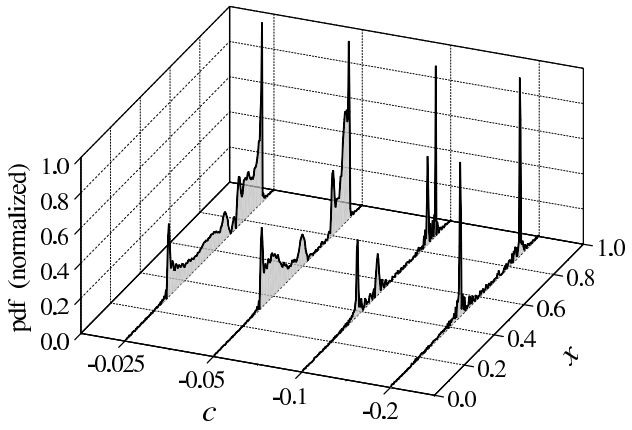

Fig. 6 Pdfs for various values of c with Algorithm 1 ($a = 3.7$ and $N = 256$).

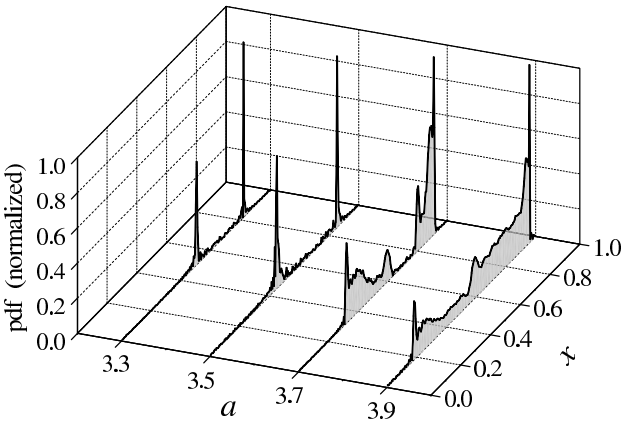

Fig. 8 Pdfs for various values of a with Algorithm 1 ($c = -0.05$ and $N = 256$).

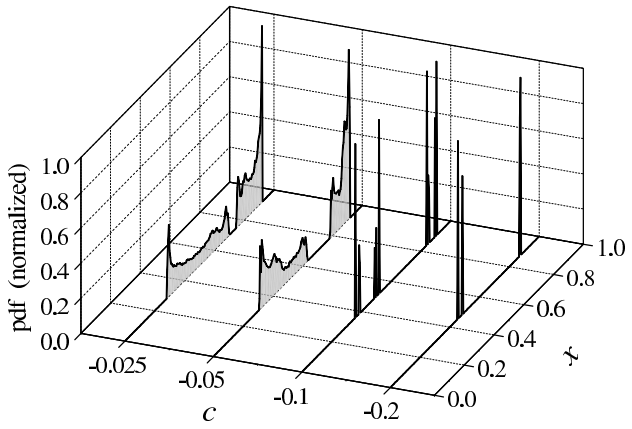

Fig. 7 Pdfs for various values of c by numerical simulation ($a = 3.7$).

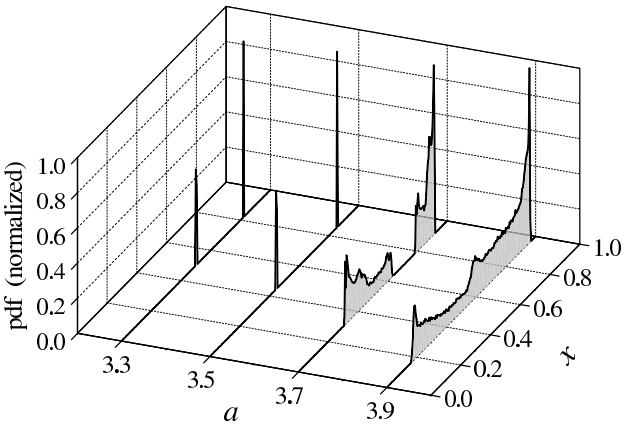

Fig. 9 Pdfs for various values of a by numerical simulation ($c = -0.05$).

numerical simulation. When $a = 3.2$ and $c = -0.2$, x_t converged to a certain value or oscillated periodically with a period of 2 in a steady state. These two solutions were obtained by numerical simulation with different initial values of x_t, as shown in Fig. 10. The initial values were distributed between $x_{0(\mathrm{low})}$ and 1.0, and the lower boundary of the initial value, $x_{0(\mathrm{low})}$, was varied. However, the range of $x_{0(\mathrm{low})}$ such that x_t converged to a certain value was very narrow. It is thus almost impossible to set the initial value for the solution if we do not know that the solution exists. In contrast, the two pdfs were easily obtained using Algorithm 2 with $T_\rho = 1$ and $T_\rho = 2$. The pdf when x_t converged was obtained by setting $T_\rho = 1$. The pdf when x_t oscillated with a period of 2 was obtained by setting $T_\rho = 2$. Here, $\boldsymbol{\rho}_{\mathrm{sol}(1)}$ was not equal to $\boldsymbol{\rho}_{\mathrm{sol}}$, so $\boldsymbol{\rho}_{\mathrm{e}1}$ and $\boldsymbol{\rho}_{\mathrm{e}2}$ were set to $\boldsymbol{\rho}_{\mathrm{sol}}$ and $\boldsymbol{\rho}_{\mathrm{sol}(1)}$, respectively, in Step (2-4). That is, a different solution from that first derived was obtained by performing Step (2-3) only once, and the two pdfs could be more easily obtained using Algorithm 2 than using numerical simulation. This means that the initial value set in Step (2-3) was appropriate for obtaining the second solution and that using the eigenvectors of $A_{\mathrm{comb}}(\boldsymbol{\rho}_{\mathrm{sol}}, T_\rho)$ as the initial values succeeded. The pdf obtained using Algorithm 2 and that using numerical simulation are shown in Fig. 11, where the former is indicated by "MVE" and the latter by "Num." on the abscissa. We can see that the pdf when x_t converged and that when x_t oscillated were obtained.

The number of elements, L, often affects numerical simulation. That is, the pdf obtained by numerical simulation often depends on L. To reduce the effect of L on the comparisons of Algorithms 1 and 2 with numerical simulation, L was set to a sufficiently large number, 128, as mentioned at the beginning of this section. Expanding Algorithms 1 and 2 to analyze the change in the pdf with L is left for future work.

Table 1 shows the number of iterations (Steps (1-3) through (1-6)) in Algorithm 1 required to derive the pdfs in Figs. 6, 8, and 11. The calculation costs nec-

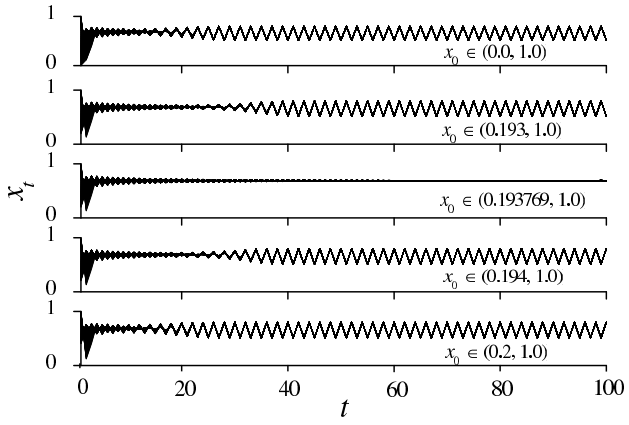

Fig. 10 Effect of initial value on dynamics in x_t ($a = 3.2$ and $c = -0.2$).

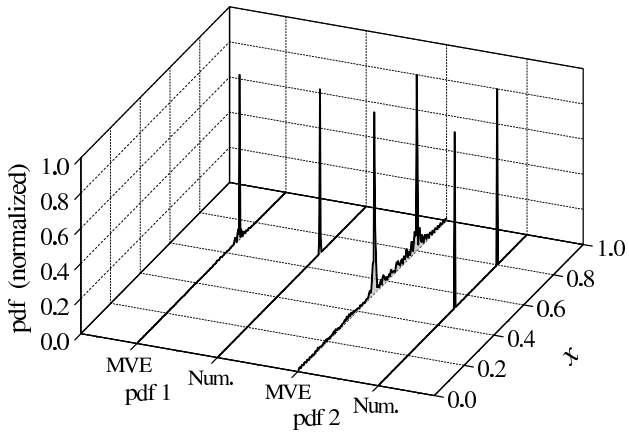

Fig. 11 Effect of initial value on pdfs of x_t ($a = 3.2$ and $c = -0.2$).

essary to execute one iteration are not small because eigen analysis is executed at each iteration. In particular, Algorithm 2 must be executed for various values of T_ρ when multiple pdfs are derived. However, the number of iterations is very small, as shown in Table 1. It is thus possible without too much calculation cost to find solutions that are difficult to do using numerical simulation. Therefore, the method presented in this paper is quite promising for analyzing many-body systems although the effect of setting the initial values on the basis of the eigenvectors has not yet been elucidated and that Algorithm 2 can always derive all the pdfs for Eq. (20) has not been proven.

5. Conclusion

A method based on a moment vector equation (MVE) was developed for analyzing many-body systems expressed by the linear interactions of identical and indistinguishable elements with nonlinear dynamics. This method has three attractive features. The first is approximation of the system by using a nonlinear MVE that avoids the curse of dimensionality. The second is algorithms that use eigen analysis of the coefficient matrix of the MVE. This analysis clarifies the structure of the solutions in the moment vector space. The third is the definite procedure of the method based on eigen analysis to effectively find multiple solutions that are difficult to do using numerical simulation. The system can be analyzed without using an intuitive or empirical procedure.

Additional work remains. The method should be enhanced so that it can be used to analyze more complex systems such as those with nonlinear interactions, and so that it can derive all the solutions. Although these are challenging tasks, this method is quite promising for analyzing many-body systems because of the attractive features, so these challenges will be addressed in a future study.

Table 1 Number of iterations required to derive pdfs in Figs. 6, 8, and 11.

Fig. 6	$c = -0.025$	$c = -0.05$	$c = -0.10$	$c = -0.20$
	2	2	2	3
Fig. 8	$a = 3.3$	$a = 3.5$	$a = 3.7$	$a = 3.9$
	2	2	2	2
Fig. 11	$T_\rho = 1$	$T_\rho = 2$		
		$\boldsymbol{\rho}_{\mathrm{sol}}$	$\boldsymbol{\rho}_{\mathrm{sol}(1)}$	–
	4	8	7	–

References

[1] K. Kaneko, "Clustering coding switching hierarchical ordering and control in a network of chaotic elements," Physica D, vol. 41, pp. 137–172, 1990.

[2] Q. Liang, "Ad hoc wireless network traffic-self-similarity and forecasting," IEEE Communications Letters, vol. 6, no. 7, pp. 297–299, Jul. 2002.

[3] T. Tsuchiya T. Konishi and N. Gouda, "Quasiequilibria in one-dimensional self-gravitating many-body systems," Physical Review E, vol. 50, no. 4, pp. 2607–2615, Oct. 1994.

[4] J. A. S. Lima R. Silva and A. R. Plastino, "Nonextensive thermostatistics and the H theorem," Physical Review Letters, vol. 86, no. 14, pp. 2983–2941, April 2001.

[5] D. Jiulin, "The nonextensive parameter and Tsallis distribution for self-gravitating systems," Europhys. Lett., vol. 67, no. 6, pp. 893–899, Sept. 2004.

[6] D. A. Dawson, "Critical dynamics and fluctuations for a mean-field model of cooperative behavior," J. Stat. Phys., vol. 31, no. 1, pp. 29–85, April 1983.

[7] M. Shiino, "Dynamical behavior of stochastic systems of infinitely many coupled nonlinear oscillators exhibiting phase transitions of mean-field type: H theorem on asymptotic approach to equilibrium and critical slowing down of order-parameter fluctuations," Phys. Rev. A, vol. 36, no. 5, pp. 2393–2412, Sept. 1987.

[8] T. D. Frank, "Short-time correlations of many-body systems described by nonlinear Fokker-Planck equations and Vlasov-Fokker-Planck equations," Physics Letters A, vol. 337, no. 3, pp. 224–234, April 2005.

[9] D. G. Luenberger, Introduction to Dynamic Systems Theory Models & Applications, John Wiley & Sons, USA, 1979.

[10] A. F. Vakakis and M. E. King, "Resonant oscillations of a weakly coupled nonlinear layered system," Acta Mech., vol. 128, no. 1-2, pp. 59-80, 1998.
[11] K. Komatsu and H. Takata, "A Formal Linearization by the Chebyshev Interpolation and Its Applications", Proceedings of the IEEE Conference on Decision and Control, Vol.1 of 4, pp.70–75, 1996.
[12] R. V. Bobryk, "Stochastic equations of the Langevin type under a weakly dependent perturbation," J. Stat. Phys., vol. 70, no. 3/4, pp. 1045–1056, Feb. 1993.
[13] G. Muscolino, G. Ricciardi, and M. Vasta, "Stationary and non-stationary probability density function for non-linear oscillators," Int. J. Non-Linear Mech., vol. 32, no. 6, pp. 1051–64, Nov. 1997.
[14] H. Satoh, "Approximation and analysis of non-linear equations in a moment vector space," IEICE Trans. Fundamentals, vol. E89-A, no. 1, pp. 270–279, Jan. 2006.
[15] I. N. Bronshtein and K. A. Semendyayev, Handbook of Mathematics, Springer-Verlag, UK, 1997.
[16] A. Papoulis, The Fourier Integral and its Applications, McGraw-Hill, USA, 1962.
[17] E. Ott, Chaos in Dynamical Systems Second ed., Cambridge University Press, UK, 2002.

Appendix A: Basis for Function Approximation

An orthonormal basis is summarized in this appendix. Let $h(\boldsymbol{k})$ be the Fourier coefficient, $\boldsymbol{k} \stackrel{\text{def}}{=} (k_1, \cdots, k_{d_\text{x}})^\text{T} \in \boldsymbol{\mathcal{Z}}$ be the index vector of the Fourier coefficient, and $\boldsymbol{\mathcal{Z}}$ be the set of $\boldsymbol{k}$ that are used for the index vectors. The Fourier series expansion for function $f(\boldsymbol{x})$ is defined by [15]

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \boldsymbol{Z}} h(\boldsymbol{k}) K(\boldsymbol{x}, \boldsymbol{k}), \tag{A·1}$$

$$h(\boldsymbol{k}) \stackrel{\text{def}}{=} \int_{\boldsymbol{\mathcal{D}}_\text{x}} f(\boldsymbol{x}) K^*(\boldsymbol{x}, \boldsymbol{k}) \text{d}\boldsymbol{x}, \tag{A·2}$$

where $\boldsymbol{x} \stackrel{\text{def}}{=} (x_1, \cdots, x_{d_\text{x}})^\text{T}$ is the state vector of dimension d_x, $\boldsymbol{\mathcal{D}}_\text{x} \stackrel{\text{def}}{=} \{\boldsymbol{x} | x_{\min d} \le x_d \le x_{\max d}, 1 \le d \le d_\text{x}\}$ is the domain of the definition of $\boldsymbol{x}$, superscript $*$ denotes a complex conjugate, $\{K(\boldsymbol{x}, \boldsymbol{k})\}$ is a multi-dimensional orthonormal basis, and K($\boldsymbol{x}$,$\boldsymbol{k}$) is defined by

$$K(\boldsymbol{x}, \boldsymbol{k}) \stackrel{\text{def}}{=} \prod_{d=1}^{d_\text{x}} K_d(x_d, k_d). \tag{A·3}$$

Here, $\{K_d(x_d, k_d)\}$ is a one-dimensional orthonormal basis.

Let $\{\psi_i(\cdot)\}$ be a basis the element of which is defined by

$$\psi_i(\boldsymbol{x}) \stackrel{\text{def}}{=} K(\boldsymbol{x}, \boldsymbol{k}), \tag{A·4}$$

where i is the index of the basis. When $\mathcal{Z}_d \stackrel{\text{def}}{=} \{0, 1, \cdots, N_d\}$ and $\boldsymbol{\mathcal{Z}}$ is given by the Cartesian product as $\boldsymbol{\mathcal{Z}} = \mathcal{Z}_1 \times \mathcal{Z}_2 \times, \cdots, \times \mathcal{Z}_{d_\text{x}}$, the relationship between $\boldsymbol{k}$ and i can be obtained using

$$i = \sum_{d=1}^{d_\text{x}} k_d \prod_{d'=d+1}^{d_\text{x}} (N_{d'} + 1), \tag{A·5}$$

where N_d is the degree of expansion of x_d. Let N be the degree of expansion of $\boldsymbol{x}$. When Eq. (A·5) holds, N is expressed by

$$N = \prod_{d=1}^{d_\text{x}} (N_d + 1) - 1, \tag{A·6}$$

where the dimension of the feature space with the basis is $N + 1$. The relationship between i and $\boldsymbol{k}$ is referred to as the index table.

The element of the orthonormal basis based on the complex Fourier series is defined as [16]

$$K_d(x_d, k_d) \stackrel{\text{def}}{=} \begin{cases} \sqrt{\dfrac{1}{D_{\text{x}d}}} \text{ for } k_d = 0 \\ \sqrt{\dfrac{1}{D_{\text{x}d}}} \exp(-\imath \dfrac{k_d+1}{2} \omega_{0d} x_d) \text{ for } k_d = 1, 3, \cdots \\ \sqrt{\dfrac{1}{D_{\text{x}d}}} \exp(\imath \dfrac{k_d}{2} \omega_{0d} x_d) \text{ for } k_d = 2, 4, \cdots \end{cases} \tag{A·7}$$

where $\imath$ denotes the imaginary unit, $\omega_{0d} \stackrel{\text{def}}{=} 2\pi / D_{\text{x}d}$, and $D_{\text{x}d} \stackrel{\text{def}}{=} x_{\max d} - x_{\min d}$.